Molding CNNs for text: non-linear, non-consecutive convolutions

Tao Lei, Regina Barzilay, and Tommi Jaakkola

Presented by Shih-Ming Wang
NLPLab, Institute of Information Science, Academia Sinica

07-5-2016

Outline

Motivation

- Deep learning & Convolution neural network (CNN) have led to success in many NLP problems
- Convolution operation is a **linear** mapping over **n-gram** vectors
- Target: **non-linear** operation over **non-consecutive** n-grams (e.g., "not that good")

Background

Tensor-based Feature Mapping

- Intuition: use product operation to remedy the insufficiency of linear operation
- Consider 2-gram (x_1, x_2) as example:

	Linear	Product
Dim(x)	$1 \times d$	$1 \times d$
Dim(filter) $h \times 2 \times d$	$h \times 2 \times d^2$	
Output	$1 \times h$	$h \times d^2$ height

Model Description

Experiments

Error Analysis